Highlights

Bayesian definition of random sequences with respect to conditional probabilities[⋆]

Hayato Takahashi

- Algorithmic randomness for conditional probabilities is studied.

- Blind randomness is ill-defined for conditional probabilities.

- Effective orthogonality and existence of consistent estimator are equivalent.

- An algorithmic solution to a classical problem in Bayes statistics.

Bayesian definition of random sequences with respect to conditional probabilities

Hayato Takahashi[a]

[a]*Random Data Lab. Inc., 3-8-18 Minami-Hanahata Adachi-ku, Tokyo, 1210062, Japan*

Abstract

We study Martin-Löf random (ML-random) points on computable probability measures on sample and parameter spaces (Bayes models). We consider variants of conditional randomness defined by ML-randomness on Bayes models and those of conditional blind randomness. We show that variants of conditional blind randomness are ill-defined from the Bayes statistical point of view. We prove that if the sets of random sequences of uniformly computable parametric models are pairwise disjoint then there is a consistent estimator for the model. Finally, we present an algorithmic solution to a classical problem in Bayes statistics, i.e. the posterior distributions converge weakly to almost all parameters if and only if the posterior distributions converge weakly to all ML-random parameters.

Keywords: Martin-Löf randomness, generalized van Lambalgen's theorem, conditional probability, collective, Bayes consistency theorem, uniform randomness
MSC[2020] 03D32, 68Q30

1. Introduction

We study Martin-Löf random [17–19, 22, 23, 28] (ML-random) points on computable probability measures on sample and parameter spaces (computable Bayes models). We assume that samples and parameters are infinite binary sequences except for Section 4.2. A conditional distribution [16, 36] is defined for a Bayes model. A conditional distribution given a finite prefix of sample sequence is called a *posterior distribution*. A marginal distribution on parameter space is called a *prior*.

In Bayes statistics, we study relations between samples and parameters. Loosely speaking, we say that a probability model is a *true model* of a sequence or a sequence is *generated* by the model if the sequence is random with respect to (w.r.t.) the model. In Bayes statistics, we assume that a sample sequence x^∞ is generated by the marginal distribution on

---

sample space. Then we estimate the parameter y^∞ by the posterior distributions given finite prefixes of x^∞ such that x^∞ is random w.r.t. the conditional distribution given y^∞. The Bayes consistency theorem [8, 11, 14] says that the posterior distributions given finite prefixes of x^∞ weakly converge to a parameter y^∞ for almost all x^∞ w.r.t. the conditional distribution given y^∞ for almost all y^∞ w.r.t. the prior for an appropriate class of Bayes models (consistent Bayes models). The parameter of the true model is not determined a priori but is estimated by the posterior distributions cf. von Mises [25]. If we consider only a fixed parameter, we do not know if the posterior distributions given finite prefixes of random sequence weakly converge to some parameter.

To state the Bayes consistency theorem for individual random sequences and parameters, we need to consider the sets of random sequences and parameters w.r.t. marginal distributions and the family of sets of random sequences w.r.t. conditional distributions given random parameters simultaneously. For computable Bayes models, we assume that (i) *a sequence is random w.r.t. the marginal distribution on sample space if and only if there is a random parameter and the sequence is random w.r.t. the conditional distribution given the parameter.* For computable and consistent Bayes models, we further assume that (ii) *if a sequence is random w.r.t. the marginal distribution on sample space then the posterior distributions given finite prefixes of the sequence weakly converge to the true model.* The assumption (i) is equivalent to that (i') *the set of random sequences w.r.t. the marginal distribution on sample space equals the union of the sets of random sequences w.r.t. the conditional distributions given random parameters.* The family of the sets of random sequences w.r.t. the conditional distributions given random parameters is determined by those distributions (and hence by the Bayes model). The assumption (i') requires that the union of those sets is determined only by the marginal distribution on sample space. Assumption (i), (i'), and (ii) are natural requirements for any notion of randomness in Bayes models, see Section 4.3.

Theorem 2.1 [30–32] shows that for a computable Bayes model there exists a version of conditional distribution (the standard conditional distribution) that is defined for ML-random parameters w.r.t. the prior. We study four families of the sets of random sequences w.r.t. the standard conditional distribution: two of them are defined by ML-randomness on Bayes models (variants of conditional Bayes randomness) and the others are defined by the blind (Hippocratic) tests [6, 15] for conditional distributions with given parameters (variants of conditional blind randomness). Theorem 2.4 parts 1 and 3 show that variants of conditional Bayes randomness satisfy the assumptions (i') and (ii) for all computable Bayes models. On the other hand, there are different computable and consistent Bayes models that have the same marginal distribution on sample space but their unions of conditional blind random sequences for all random parameters are different. Theorem 2.4 part 4 shows that there is no family of sets of random sequences w.r.t. the marginal distribution on sample space that satisfy the assumption (i') for variants of conditional blind randomness for the class of computable and consistent Bayes models. We consider that variants of conditional blind randomness are ill-defined as conditional randomness for Bayes models.

The rest of the paper is structured as follows. In Section 2, we state our main theorem. In Section 3, we study ML-random points in Bayes models. In Section 3.1, we show the rela-

2

tionship between variants of random sequences w.r.t. the standard conditional distributions. In this paper, we assume the computability of joint probability measures, however, we do not demand uniform computability on standard conditional distributions. In Section 4.1, we compare ML-randomness in Bayes models with randomness for uniformly computable parametric models [6, 13, 15, 21, 35]. Theorem 4.6 shows that standard conditional distributions are equal to uniformly computable parametric models on ML-random parameters when the Bayes models are constructed from uniformly computable parametric models and computable priors. We show that if the sets of uniform random sequences are pairwise disjoint for different parameters (effectively orthogonal [6]), then there is a consistent estimator, i.e. a measurable function that equals y^∞ with probability one for the probability model of the parameter y^∞ for all y^∞. In Section 4.2, we discuss ML-randomness of Bayes models on complete separable metric spaces and present an algorithmic solution to a classical problem in Bayes statistics [9], i.e. for the Bayes models on complete separable metric spaces, the posterior distribution is consistent at almost all parameters if and only if the posterior distribution is consistent at all ML-random parameters. Finally in Section 4.3, we briefly re-discuss randomness in statistical models. We compare our notion of randomness in Bayes models with that in non-Bayes models (parametric models without prior), which may help to understand our assumptions (i), (i'), and (ii).

2. Main theorem

Let Ω be the set of infinite binary sequences, and S the set of finite binary strings. We also call an element of $\Omega \cup \Omega^2$ a point. Let $\Delta(x)$ be the set of infinite binary sequences that start with $x \in S$, and $|x|$ the length of $x \in S$. To clarify the difference between finite strings and infinite sequences, we use symbols such as x, y for finite strings and x^∞, y^∞ for infinite sequences. Except for the clearly stated cases, do not confuse symbols such as x^∞ with the repetition of a string. Let λ be the empty word. We write $x \sqsubseteq y$ if x is a prefix of y for $x, y \in S \cup \Omega$ including the case $x = y$ and $x \sqsubset y$ if $x \sqsubseteq y$ and $x \neq y$. Let B^c be the complement of B. We write $\tilde{A} := \bigcup_{x \in A} \Delta(x)$ for $A \subseteq S$. For $A \subseteq S^2$, $\tilde{A}$ is defined similarly.

Let open sets in Ω be those generated by $\{\Delta(x) \mid x \in S\}$, i.e. every open set in Ω is a union of elements in $\{\Delta(x) \mid x \in S\}$. Let $(\Omega, \mathcal{B}_1)$ be a measurable space where $\mathcal{B}_1$ is the smallest σ-algebra that includes $\{\Delta(x) \mid x \in S\}$. Similarly, let open sets in Ω^2 be those generated by $\{\Delta(x) \times \Delta(y) \mid x, y \in S\}$ and $(\Omega^2, \mathcal{B}_2)$ a measurable space where $\mathcal{B}_2$ is the smallest σ-algebra that includes $\{\Delta(x) \times \Delta(y) \mid x, y \in S\}$. Let $d(x^\infty, y^\infty) := \sum_n |x_n - y_n| 2^{-n}$ for all $x^\infty, y^\infty \in \Omega$ where $x^\infty = x_1 x_2 \cdots$ and $y^\infty = y_1 y_2 \cdots$. Then d is a metric on Ω and open sets induced by d are equivalent to those generated by $\{\Delta(x) \mid x \in S\}$. Similarly for $(x^\infty, y^\infty), (x'^\infty, y'^\infty) \in \Omega^2$, let $d^2((x^\infty, y^\infty), (x'^\infty, y'^\infty)) := d(x^\infty, x'^\infty) + d(y^\infty, y'^\infty)$. Then d^2 is a metric on Ω^2 and open sets induced by d^2 are equivalent to those generated by $\{\Delta(x) \times \Delta(y) \mid x, y \in S\}$. In the following, we consider the metrics d on Ω and d^2 on Ω^2.

Except for Section 4.2, set $X = Y = \Omega$. Let P be a probability measure on $(X \times Y, \mathcal{B}_2)$, P_X marginal distribution on $(X, \mathcal{B}_1)$, and P_Y marginal distribution on $(Y, \mathcal{B}_1)$. For all $x, y \in S$, let $P(x, y) := P(\Delta(x) \times \Delta(y))$, $P_X(x) := P(\Delta(x) \times \Omega)$, and $P_Y(y) := P(\Omega \times \Delta(y))$. For

3

all $x, y \in S$, let $P(x \mid y) := P(x, y)/P_Y(y)$ if $P_Y(y) \neq 0$ and $P_{Y|X}(y \mid x) := P(x, y)/P_X(x)$ (posterior distribution) if $P_X(x) \neq 0$.

Let $U \subseteq S \times \mathbb{N}$. U is called a *test* (effective null set) or ML-test w.r.t. P on $(\Omega, \mathcal{B})$ if U is recursively enumerable (r.e.), $\tilde{U}_n \supseteq \tilde{U}_{n+1}$, and $P(\tilde{U}_n) < 2^{-n}$, where $U_n = \{x \mid (x, n) \in U\}$, for all n. The ML-random sequences w.r.t. P are defined as the complement of the effective null sets w.r.t. P. We denote it by $\mathcal{R}^P$, i.e. $\mathcal{R}^P := (\bigcup_{U : \, test} \bigcap_n \tilde{U}_n)^c$. Let $\mathcal{R}^{P,A} := (\bigcup_{U : \, test} \bigcap_n \tilde{U}_n)^c$, where U is a test with oracle A, i.e. U is r.e. with oracle A, $\tilde{U}_n \supseteq \tilde{U}_{n+1}$, and $P(\tilde{U}_n) < 2^{-n}$ for all n. Similarly, $\mathcal{R}^P$ and $\mathcal{R}^{P,A}$ are defined w.r.t. P on $(\Omega^2, \mathcal{B}_2)$. $\mathcal{R}^P$ and $\mathcal{R}^{P,A}$ are forms of blind randomness, i.e. we neither assume that P is computable nor computable with oracle A. In the following, for simplicity, we say that $(\Omega, \mathcal{B}_1, P)$ and $(X \times Y, \mathcal{B}_2, P)$ are computable if P on $(\Omega, \mathcal{B}_1)$ and $(\Omega^2, \mathcal{B}_2)$ are computable, respectively.

We obtain the following theorem from the martingale convergence theorem for individual ML-random sequences [30, 31].

Theorem 2.1 (Takahashi [30–32]) *Assume that $(X \times Y, \mathcal{B}_2, P)$ is computable. For all $x \in S$ and $y^\infty \in \mathcal{R}^{P_Y}$, set*

$$P(x \mid y^\infty) := \lim_{y \to y^\infty} P(x \mid y)$$

if the right-hand side exists. Then for each $y^\infty \in \mathcal{R}^{P_Y}$, $P(\cdot \mid y^\infty)$ is a probability measure on $(X, \mathcal{B}_1)$.

Definition 2.2 *The family of probability measures $\{P(\cdot \mid y^\infty) \mid y^\infty \in \mathcal{R}^{P_Y}\}$ is called the standard conditional distribution.*

For each $x \in S$, conditional probability $P(x \mid y^\infty)$ is a Borel-measurable random variable on Y that satisfies

$$\int_{\Delta(y)} P(x \mid y^\infty) dP_Y = P(x \mid y) \text{ for all } y \in S,$$

see Theorem 33.1 on page pp.434 in [7]. Conditional probability is not unique. Two versions of conditional probability are called equivalent if they coincide for almost all y^∞. From the Radon-Nikodým theorem, the standard conditional distribution is a version of conditional probability [36]. A version of conditional probability is called *regular* if the conditional probability given the parameter is a probability measure for almost all parameters. By Theorem 2.1, the standard conditional distribution is a probability measure for each ML-random parameter and hence regular, but it may not be computable with oracle access to the ML-random parameter, see Remark 3.3. To state the theorems in the Bayes models for individual ML-random points, e.g. consistency theorem for the posterior distributions (Theorem 3.4), we fix the standard conditional distribution as a version of conditional probability. A computable Bayes model defines a standard conditional distribution and prior, and vice versa.

We define consistent Bayes models.

$$\bigcap_{y\to y^\infty} \mathcal{R}^{P(\cdot|y)} \qquad\qquad \mathcal{R}^P_{y^\infty} = \bigcap_{y\to y^\infty} \mathcal{R}^{P(\cdot|y)}$$

(i) $\subseteq\!\!\!/$ $\searrow$ (ii) $\cup|$ (v)

$$\mathcal{R}^P_{y^\infty} \qquad \mathcal{R}^{P(\cdot|y^\infty)} \qquad \mathcal{R}^{P(\cdot|y^\infty)}$$

(iii) $\searrow$ $\subseteq\!\!\!/$ (iv) $\cup|$ (vi)

$$\mathcal{R}^{P(\cdot|y^\infty),y^\infty} \qquad\qquad \mathcal{R}^{P(\cdot|y^\infty),y^\infty}$$
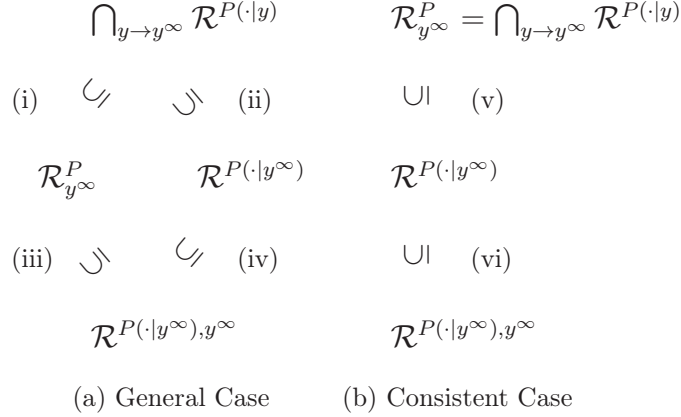
(a) General Case (b) Consistent Case

Figure 1: Figure (a) shows relations between variants of conditional random sequences. Figure (b) shows those relations when the Bayes model is consistent. If the conditional probability is computable with ML-random oracle y^∞ and the Bayes model is consistent then the four sets in figure (b) are equal. Theorem 3.10 shows a counter-example of the equalities in (ii) and (v). Lemma 3.9 shows a counter-example of the equalities in (i) and (iv). Lemma 3.9 and Theorem 3.10 show that $\mathcal{R}^P_{y^\infty}$ and $\mathcal{R}^{P(\cdot|y^\infty)}$ are incomparable in the general case. The counter-example of equality in (iii) is due to Bauwens [4, 5]. The equality in (vi) for consistent parametric models was presented in Kjos Hanssen [15] and Bienvenu et al. [6], see Section 4.1. An example of the equality in (vi) for non-uniformly computable orthogonal conditional distributions is presented in Theorem 3.10. It remains open whether the equality in (vi) always holds for all ML-random y^∞ (Remark 3.11).

Definition 2.3 *Let P be a probability measure on $(X \times Y, \mathcal{B}_2)$. Let δ_{y^∞} be the probability measure on $(Y, \mathcal{B}_1)$ such that $\delta_{y^\infty}(\{y^\infty\}) = 1$. The posterior distribution $P_{Y|X}(\cdot \mid x)$ weakly converges to δ_{y^∞} as $x \to x^\infty$ if and only if for any neighborhood A of y^∞, $\lim_{x\to x^\infty} P_{Y|X}(A \mid x) = 1$. The posterior distribution is called consistent at y^∞ if it weakly converges to δ_{y^∞} as $x \to x^\infty$ for almost all x^∞ w.r.t. $P(\cdot \mid y^\infty)$. P is called consistent if the posterior distribution is consistent at almost all parameters.*

For a set $A \subseteq \Omega \times \Omega$, let $A_{y^\infty} := \{x^\infty \mid (x^\infty, y^\infty) \in A\}$. Consider the following sets for $y^\infty \in \mathcal{R}^{P_Y}$,
(i) $\bigcap_{y\to y^\infty} \mathcal{R}^{P(\cdot|y)}$, the intersection of the sets of ML-random sequences w.r.t. $P(\cdot \mid y)$ for all finite prefixes of y^∞,
(ii) $\mathcal{R}^P_{y^\infty}$, the section of ML-random points $\mathcal{R}^P$ at y^∞,
(iii) $\mathcal{R}^{P(\cdot|y^\infty)}$, ML-random sequences w.r.t. $P(\cdot \mid y^\infty)$, and
(iv) $\mathcal{R}^{P(\cdot|y^\infty),y^\infty}$, ML-random sequences w.r.t. $P(\cdot \mid y^\infty)$ with oracle y^∞.
Figure 1 shows relations between these sets.

Theorem 2.4 (Main theorem)
1. Assume that $(X \times Y, \mathcal{B}_2, P)$ is computable. Then

$$\mathcal{R}^{P_X} = \bigcup_{y^\infty \in \mathcal{R}^{P_Y}} \bigcap_{y\to y^\infty} \mathcal{R}^{P(\cdot|y)} = \bigcup_{y^\infty \in \mathcal{R}^{P_Y}} \mathcal{R}^P_{y^\infty} \supseteq \bigcup_{y^\infty \in \mathcal{R}^{P_Y}} \mathcal{R}^{P(\cdot|y^\infty)} \supseteq \bigcup_{y^\infty \in \mathcal{R}^{P_Y}} \mathcal{R}^{P(\cdot|y^\infty),y^\infty}.$$

5

2. Assume that $(X \times Y, \mathcal{B}_2, P)$ is computable and $P(\cdot \mid y^\infty)$ is computable with oracle y^∞ for each fixed $y^\infty \in \mathcal{R}^{P_Y}$. Then

$$\mathcal{R}^{P_X} = \bigcup_{y^\infty \in \mathcal{R}^{P_Y}} \bigcap_{y \to y^\infty} \mathcal{R}^{P(\cdot|y)} = \bigcup_{y^\infty \in \mathcal{R}^{P_Y}} \mathcal{R}^P_{y^\infty} = \bigcup_{y^\infty \in \mathcal{R}^{P_Y}} \mathcal{R}^{P(\cdot|y^\infty)} = \bigcup_{y^\infty \in \mathcal{R}^{P_Y}} \mathcal{R}^{P(\cdot|y^\infty),y^\infty}.$$

3. Assume that $(X \times Y, \mathcal{B}_2, P)$ is computable and consistent. Then

$$x^\infty \in \mathcal{R}^{P_X} \text{ if and only if there is a unique } y^\infty \in \mathcal{R}^{P_Y} \text{ such that}$$
$$P_{Y|X}(x) \text{ weakly converges to } \delta_{y^\infty} \text{ as } x \to x^\infty \text{and}$$
$$x^\infty \in \mathcal{R}^P_{y^\infty} (= \bigcap_{y \to y^\infty} \mathcal{R}^{P(\cdot|y)}).$$

4. There is no $\tilde{\mathcal{R}}\colon \{P \mid (\Omega, \mathcal{B}_1, P) \text{ is computable}\} \to \{A \mid A \subseteq \Omega\}$ such that

$$\tilde{\mathcal{R}}(P_X) = \bigcup_{y^\infty \in \mathcal{R}^{P_Y}} \mathcal{R}^{P(\cdot|y^\infty)} \text{ for all computable and consistent } (X \times Y, \mathcal{B}_2, P).$$

Similarly, there is no $\tilde{\mathcal{R}}\colon \{P \mid (\Omega, \mathcal{B}_1, P) \text{ is computable}\} \to \{A \mid A \subseteq \Omega\}$ such that

$$\tilde{\mathcal{R}}(P_X) = \bigcup_{y^\infty \in \mathcal{R}^{P_Y}} \mathcal{R}^{P(\cdot|y^\infty),y^\infty} \text{ for all computable and consistent } (X \times Y, \mathcal{B}_2, P).$$

5. Assume that $(X \times Y, \mathcal{B}_2, P)$ is computable. Then
the sets (i)–(iv) are Borel sets in Ω and the probabilities of these sets w.r.t. $P(\cdot \mid y^\infty)$ are one for all $y^\infty \in \mathcal{R}^{P_Y}$, and
$\bigcup_{y^\infty \in \mathcal{R}^{P_Y}} \bigcap_{y \to y^\infty} \mathcal{R}^{P(\cdot|y)}$ and $\bigcup_{y^\infty \in \mathcal{R}^{P_Y}} \mathcal{R}^P_{y^\infty}$ are Borel sets in Ω and the probabilities of these sets w.r.t. P_X are one.

The proof of Theorem 2.4 is given after the proof of Theorem 3.10.

Roughly speaking, the sets (i) and (ii) are defined from Bayes models; those (iii) and (iv) are defined from conditional probabilities with given parameters. For simplicity, we call the families of sets $\{\bigcap_{y \to y^\infty} \mathcal{R}^{P(\cdot|y)} \mid y^\infty \in \mathcal{R}^{P_Y}\}$ and $\{\mathcal{R}^P_{y^\infty} \mid y^\infty \in \mathcal{R}^{P_Y}\}$ *variants of conditional Bayes randomness*. We call the families of sets $\{\mathcal{R}^{P(\cdot|y^\infty)} \mid y^\infty \in \mathcal{R}^{P_Y}\}$ and $\{\mathcal{R}^{P(\cdot|y^\infty),y^\infty} \mid y^\infty \in \mathcal{R}^{P_Y}\}$ *variants of conditional blind randomness*. By Theorem 2.4 parts 1 and 3, both of the variants of conditional Bayes randomness satisfy our assumptions (i') and (ii). On the other hand, by Theorem 2.4 part 4, neither variant of conditional blind randomness satisfies the assumption (i'). We consider that variants of conditional blind randomness are ill-defined as conditional randomness for Bayes models.

Takahashi [30–33] defined the family of sets $\{\mathcal{R}^P_{y^\infty} \mid y^\infty \in \mathcal{R}^{P_Y}\}$ to be the sets of conditional random sequences w.r.t. the standard conditional distribution $\{P(\cdot \mid y^\infty) \mid y^\infty \in \mathcal{R}^{P_Y}\}$.

Remark 2.5 The author does not know if the sets $\bigcup_{y^\infty \in \mathcal{R}^{P_Y}} \mathcal{R}^{P(\cdot|y^\infty)}$ and $\bigcup_{y^\infty \in \mathcal{R}^{P_Y}} \mathcal{R}^{P(\cdot|y^\infty),y^\infty}$ are always measurable for all computable $(X \times Y, \mathcal{B}_2, P)$.

6

3. ML-randomness and Bayes models

The van Lambalgen theorem [20] states that a pair of sequences $(x^\infty, y^\infty) \in \Omega^2$ is ML-random w.r.t. the product of uniform measures if and only if y^∞ is ML-random and x^∞ is ML-random with oracle y^∞.

Theorem 3.1 (Generalized van Lambalgen theorem, Takahashi [30–33]) *Assume that* $(X \times Y, \mathcal{B}_2, P)$ *is computable. Then*

$$\mathcal{R}_{y^\infty}^P \supseteq \mathcal{R}^{P(\cdot|y^\infty), y^\infty} \text{ for all } y^\infty \in \mathcal{R}^{P_Y}. \tag{1}$$

Fix $y^\infty \in \mathcal{R}^{P_Y}$ *and assume that* $P(\cdot \mid y^\infty)$ *is computable with oracle* y^∞. *Then*

$$\mathcal{R}_{y^\infty}^P = \mathcal{R}^{P(\cdot|y^\infty), y^\infty}. \tag{2}$$

Theorem 3.2 (Takahashi [30–32]) *Assume that* $(X \times Y, \mathcal{B}_2, P)$ *is computable. Then*

$$P(\mathcal{R}_{y^\infty}^P \mid y^\infty) = 1 \text{ if } y^\infty \in \mathcal{R}^{P_Y} \text{ and } \mathcal{R}_{y^\infty}^P = \emptyset \text{ else.}$$

$$\mathcal{R}^{P_X} = \bigcup_{y^\infty \in \mathcal{R}^{P_Y}} \mathcal{R}_{y^\infty}^P. \tag{3}$$

Remark 3.3 1. Vovk and V'yugin [35] proved (2) and (3) for Bayes models that are constructed from uniformly computable parametric models and computable priors, see Section 4.1.
2. Theorem 5.2 in [31] demonstrates equation (1) when $P(\cdot \mid y^\infty)$ is computable with oracle y^∞. However, the same proof for (1) holds true when $P(\cdot \mid y^\infty)$ is not computable with oracle y^∞. In [33], (2) is proved without assuming uniform computability of the conditional distributions.
3. Non-computable conditional distributions are presented in work by Ackerman et al. [1]. Bauwens [4] showed an example that violates the equality in (2) and $\mathcal{R}^{P(\cdot|y^\infty), y^\infty}$ is a proper subset of $\mathcal{R}_{y^\infty}^P$ in (1) when the conditional distribution is not computable with oracle y^∞. Takahashi [34] showed an example that the conditional distributions are not computable with oracle y^∞ for all y^∞, but (2) holds true. For more details on the generalized van Lambalgen theorem, see the survey Bauwens et al. [5]. □

Two probability measures P and Q on $(\Omega, \mathcal{B}_1)$ are called *orthogonal* and denoted by $P \perp Q$ if there is $A \in \mathcal{B}_1$ such that $P(A) = 1$ and $Q(A) = 0$. The following theorem shows equivalent statements for the consistency of Bayes models including statements described with ML-randomness.

Theorem 3.4 *Assume that* $(X \times Y, \mathcal{B}_2, P)$ *is computable. The following statements are equivalent:*
(i) $P(\cdot \mid y) \perp P(\cdot \mid z)$ *for all* $\Delta(y) \cap \Delta(z) = \emptyset$.
(ii) $\mathcal{R}^{P(\cdot|y)} \cap \mathcal{R}^{P(\cdot|z)} = \emptyset$ *for all* $\Delta(y) \cap \Delta(z) = \emptyset$.

(iii) $P_{Y|X}(\cdot \mid x)$ weakly converges to δ_{y^∞} as $x \to x^\infty$ for all $(x^\infty, y^\infty) \in \mathcal{R}^P$.

(iv) $\mathcal{R}^{P_Y} \subseteq \{y^\infty \in \Omega \mid$ the posterior distribution $P_{Y|X}$ is consistent at $y^\infty\}$.

(v) P is consistent.

(vi) There is a measurable $f \colon X \to Y$ such that $P(f^{-1}(y^\infty) \mid y^\infty) = 1$ for almost all y^∞ w.r.t. P_Y.

(vii) $\mathcal{R}^P_{y^\infty} \cap \mathcal{R}^P_{z^\infty} = \emptyset$ for all $y^\infty \neq z^\infty$.

(viii) There is a measurable onto $f \colon \mathcal{R}^{P_X} \to \mathcal{R}^{P_Y}$ such that $P(f^{-1}(y^\infty) \mid y^\infty) = 1$ for all $y^\infty \in \mathcal{R}^{P_Y}$.

Before we prove the theorem, we show a lemma.

Lemma 3.5 *Assume that* $(X \times Y, \mathcal{B}_2, P)$ *is consistent. Then there is a measurable* $f \colon X \to Y$ *such that* $P(f^{-1}(y^\infty) \mid y^\infty) = 1$ *for almost all* y^∞ *w.r.t.* P_Y.

Proof) Let

$$g^{-1}(y^\infty) := \{x^\infty \mid P_{Y|X}(\cdot \mid x) \text{ weakly converges to } \delta_{y^\infty} \text{ as } x \to x^\infty\} \text{ and}$$

$$f^{-1}(y^\infty) := \begin{cases} g^{-1}(y^\infty) & \text{if } y^\infty \neq 0^\infty \\ g^{-1}(0^\infty) \cup (X \setminus \bigcup_{y^\infty \in Y} g^{-1}(y^\infty)) & \text{else,} \end{cases}$$

where 0^∞ is the sequence consisting of 0s.

Then $f \colon X \to Y$ is well-defined, i.e. $f^{-1}(y^\infty) \cap f^{-1}(z^\infty) = \emptyset$ for all $y^\infty \neq z^\infty$. We show that f is measurable. Let

$$A_{y,\epsilon,L} := \{x \mid P_{Y|X}(y \mid x) > 1 - \epsilon, \ |x| = L\} \text{ and}$$

$$B_y := \bigcap_{k \colon \epsilon = 1/k} \bigcup_L \bigcap_{L' \colon L \leq L'} \tilde{A}_{y,\epsilon,L'}.$$

We have that $B_y = \{x^\infty \mid \lim_{x \to x^\infty} P_{Y|X}(y \mid x) = 1\}$, $B_y \cap B_w = \emptyset$ if $\Delta(y) \cap \Delta(w) = \emptyset$, and $B_y \supseteq B_w$ if $y \sqsubseteq w$. By the definition of weak convergence of the posterior distributions, we have

$$\bigcap_{n \colon |z| \leq n} \bigcup_{\substack{y \colon |y| = n \\ z \sqsubseteq y}} B_y = g^{-1}(\Delta(z)).$$

Since $\tilde{A}_{y,\epsilon,L}$ is open for all y, ϵ, L, we have that $g^{-1}(\Delta(z)) \in \mathcal{B}_1$ for all $z \in S$ and f is measurable. Since P is consistent, we have $P(f^{-1}(y^\infty) \mid y^\infty) = 1$ for almost all y^∞. $\square$

Proof of Theorem 3.4) For the proof of (i) $\Rightarrow$ (ii) $\Rightarrow$ (iii), see the proof of Theorem 6.1 in [33]. The implication (iii) $\Rightarrow$ (iv) is immediate. The implication (iv) $\Rightarrow$ (v) follows from that $P_Y(\mathcal{R}^{P_Y}) = 1$. The implication (v) $\Rightarrow$ (vi) is due to Lemma 3.5. We show (vi) $\Rightarrow$ (i). By (vi), we have $f^{-1}(\Delta(y)), f^{-1}(\Delta(z)) \in \mathcal{B}_1$ and $f^{-1}(\Delta(y)) \cap f^{-1}(\Delta(z)) = \emptyset$ for all $\Delta(y) \cap \Delta(z) = \emptyset$. By the Fubini theorem and (vi), we have $P(f^{-1}(\Delta(y)) \mid y) = 1$ and $P(f^{-1}(\Delta(z)) \mid z) = 1$ for all $\Delta(y) \cap \Delta(z) = \emptyset$. We have the equivalence (i)–(vi).

The implication (iii) $\Rightarrow$ (vii) is immediate. We show (vii) $\Rightarrow$ (viii). Let $f^{-1}(y^\infty) = \mathcal{R}^P_{y^\infty}$ for all $y^\infty \in \mathcal{R}^{P_Y}$. By (vii) and Theorem 3.2, $f \colon \mathcal{R}^{P_X} \to \mathcal{R}^{P_Y}$ is well-defined, an onto function,

and $P(f^{-1}(y^\infty) \mid y^\infty) = 1$ for all $y^\infty \in \mathcal{R}^{P_Y}$. By Theorem 3.2, $f^{-1}(\mathcal{R}^{P_Y}) = \mathcal{R}^{P_X} \in \mathcal{B}_1$ and similarly, $f^{-1}(\Delta(s) \cap \mathcal{R}^{P_Y}) \in \mathcal{B}_1$ for any $s \in S$, and we see that f in (viii) is measurable. The proof of (viii) $\Rightarrow$ (i) is similar to that of (vi) $\Rightarrow$ (i). $\qquad\square$

Remark 3.6 Takahashi [30, 32, 33] showed the equivalence of the statement in Theorem 3.4 except for statements (iv) and (v). The equivalence (i) and (vi) is due to Corollary 2 in Breiman et al. [8]. The equivalence (i) and (ii) is due to Martin-Löf (pp. 103 second paragraph [24]), see also Theorem 4.1 in [33] and Theorem 8.6 in [6]. Doob [11] showed that if there is a measurable function $f: X \to Y$ such that $P(f^{-1}(y^\infty) \mid y^\infty) = 1$ for all $y^\infty \in Y$ then P is consistent, see also [14, 27].

For further study on ML-random points on Bayes models, see Dębowski [10], Takahashi [31, 33] and Vovk and V'yugin [35]. When the prior is discrete, Theorem 3.4 reduces to the consistency of the MDL model selection [2, 3] for individual random sequences [33]. Li and Vitányi [22] study MDL model selection in terms of Kolmogorov complexity.

3.1. Relations between conditional randomness
    We prove all inclusions shown in Figure 1.

Lemma 3.7 *Assume that $(X \times Y, \mathcal{B}_2, P)$ is computable. Then*

$$\mathcal{R}^{P(\cdot|y^\infty)} \subseteq \mathcal{R}^{P(\cdot|y)} \text{ for all } y \sqsubset y^\infty \text{ and } y^\infty \in \mathcal{R}^{P_Y}.$$

Proof) First we prove the lemma for $y = \lambda$. Let $y = \lambda$ then $P(\cdot \mid y) = P_X$. Let U^X be a test w.r.t. P_X and $U^{X\times\lambda} := \{(x, \lambda, n) \mid (x, n) \in U^X\}$. $U^{X\times\lambda}$ is a test w.r.t. P. Since $(\tilde{U}_n^{X\times\lambda})_{y^\infty} = \tilde{U}_n^X$, from Corollary 4.1 in [31], for all $y^\infty \in \mathcal{R}^{P_Y}$ there is an integer M such that

$$P(\sum_n I_{\tilde{U}_n^X} > k \mid y^\infty) < \frac{M}{k} \text{ for all } k,$$

where I is the characteristic function, i.e. $I_{\tilde{U}_n^X}(x^\infty) = 1$ if $x^\infty \in \tilde{U}_n^X$ and 0 otherwise. We have a test V^X ($\tilde{V}_k^X := \{\sum_n I_{\tilde{U}_n^X} > M2^k\}$ for all k) w.r.t. $P(\cdot \mid y^\infty)$ and the lemma is proved for $y = \lambda$. Similarly, we can show the lemma for all finite prefix $y \sqsubset y^\infty$. $\qquad\square$

Theorem 3.8 *Assume that $(X \times Y, \mathcal{B}_2, P)$ is computable.*

$$\bigcap_{y \to y^\infty} \mathcal{R}^{P(\cdot|y)} \supseteq \mathcal{R}^{P(\cdot|y^\infty)} \text{ and } \bigcap_{y \to y^\infty} \mathcal{R}^{P(\cdot|y)} \supseteq \mathcal{R}_{y^\infty}^P \text{ for all } y^\infty \in \mathcal{R}^{P_Y}. \qquad (4)$$

If P is consistent, we have

$$\bigcap_{y \to y^\infty} \mathcal{R}^{P(\cdot|y)} = \mathcal{R}_{y^\infty}^P \supseteq \mathcal{R}^{P(\cdot|y^\infty)} \supseteq \mathcal{R}^{P(\cdot|y^\infty), y^\infty} \text{ for all } y^\infty \in \mathcal{R}^{P_Y}. \qquad (5)$$

Fix $y^\infty \in \mathcal{R}^{P_Y}$ and assume that $P(\cdot \mid y^\infty)$ is computable with oracle y^∞. If P is consistent, we have

$$\bigcap_{y \to y^\infty} \mathcal{R}^{P(\cdot|y)} = \mathcal{R}_{y^\infty}^P = \mathcal{R}^{P(\cdot|y^\infty)} = \mathcal{R}^{P(\cdot|y^\infty), y^\infty}. \qquad (6)$$

9

Proof) Similarly, as presented in Theorem 3.2, we can show that

$$\mathcal{R}^{P(\cdot|y)} = \bigcup_{y^\infty \in \mathcal{R}^{P_Y} \cap \Delta(y)} \mathcal{R}^P_{y^\infty} \text{ for all } y. \tag{7}$$

From Lemma 3.7, we obtain (4).

Assume that P is consistent. By Theorem 3.4, we have $\mathcal{R}^P_{y^\infty} \cap \mathcal{R}^P_{z^\infty} = \emptyset$ if $y^\infty \neq z^\infty$. From (7), we have

$$\bigcap_{y \to y^\infty} \mathcal{R}^{P(\cdot|y)} = \mathcal{R}^P_{y^\infty}. \tag{8}$$

From (4) and (8) we have (5). Eq. (6) follows from Theorem 3.1 and (5). □

The following lemma shows a counterexample that the inclusions (i) and (iv) in Figure 1 can be strict simultaneously.

Lemma 3.9 *Let $P := U \times U$, where U is the uniform probability measure on $(\Omega, \mathcal{B}_1)$, then $\bigcap_{y \to y^\infty} \mathcal{R}^{P(\cdot|y)} \setminus \mathcal{R}^P_{y^\infty} \neq \emptyset$ and $\mathcal{R}^{P(\cdot|y^\infty)} \setminus \mathcal{R}^{P(\cdot|y^\infty), y^\infty} \neq \emptyset$ for all $y^\infty \in \mathcal{R}^U$.*

Proof) The diagonal set is covered by a test, i.e. $\{(x^\infty, x^\infty) \mid x^\infty \in \Omega\} \subseteq (\mathcal{R}^P)^c$. Let $y^\infty \in \mathcal{R}^U$, from van Lambalgen's theorem [20], we have $y^\infty \notin \mathcal{R}^{U, y^\infty} = \mathcal{R}^P_{y^\infty} \subseteq \mathcal{R}^U = \bigcap_{y \to y^\infty} \mathcal{R}^{P(\cdot|y)}$. □

To show an example in which $\mathcal{R}^P_{y^\infty} \setminus \mathcal{R}^{P(\cdot|y^\infty)} \neq \emptyset$ in (5), we construct a consistent Bayes model with the properties (i)–(iv) listed in Theorem 3.10 by modifying the examples of Bauwens [4] (Example 2 in [5]). Note that the examples in [4, 5] do not imply (iv) in Theorem 3.10.
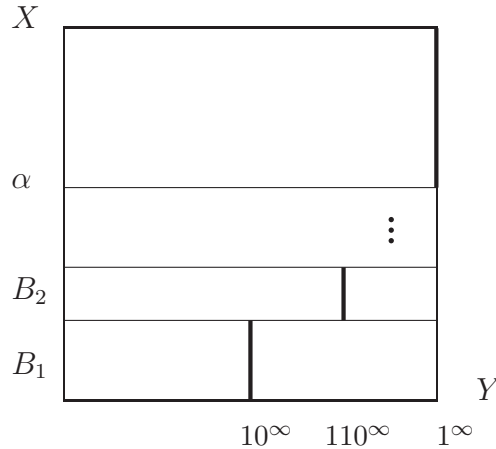

Figure 2: Construction of the counter-example. The joint measure P concentrates on the thick lines. P is consistent. The conditional distribution $P(\cdot \mid y^\infty)$ is not continuous at the parameter $y^\infty = 1^\infty$.

Theorem 3.10 *There is $(X \times Y, \mathcal{B}_2, P)$ such that*
(i) P is computable,

(ii) $\{1^\infty\} \in \mathcal{R}^{P_Y}$ and $P(\cdot \mid 1^\infty)$ is not computable,
(iii) $P(\cdot \mid y) \perp P(\cdot \mid z)$ for all $\Delta(y) \cap \Delta(z) = \emptyset$, and
(iv) $\mathcal{R}^P_{1\infty} \setminus \mathcal{R}^{P(\cdot \mid 1^\infty)} \neq \emptyset$ and $\mathcal{R}^{P_X} \setminus \bigcup_{y^\infty \in \mathcal{R}^{P_Y}} \mathcal{R}^{P(\cdot \mid y^\infty)} \neq \emptyset$.

Here, 1^∞ is the sequence consisting of 1s.

Proof of Theorem 3.10) First, we prove (i). Let P_X be the uniform distribution on X. Let $\alpha' = \sum_{s \in S} 2^{-K(s)}$, where K is the prefix complexity. Let $\alpha \in \Omega$ be the binary expansion of α' then $\alpha \in \mathcal{R}^{P_X}$ [22]. Let $g(s) := \sum_{1 \leq i \leq k} 2^{-i} s_i$ where $s = s_1 \cdots s_k$ and $s_i = 0$ or 1 for $i = 1, \ldots, k$. Consider a computable sequence of strings $a_1, a_2, \ldots$ such that $g(a_1) < g(a_2) < \cdots$ is increasing and $\lim_i g(a_i) = \alpha'$.

For all i, set $A_i := \{s \in S \mid g(s) \leq g(a_i)\}$, $B_i := \tilde{A}_i \setminus \tilde{A}_{i-1}$ for $i \geq 1$, and $A_0 = \emptyset$. We define the measure P on $X \times Y$ as follows. For all $x \in S$ and i, define

$$P(\lambda, 0) := 0,$$
$$P((B_i \cap \Delta(x)) \times \{1^i 0^\infty\}) := P_X(B_i \cap \Delta(x)),$$
$$P(B_i^c \times \{1^i 0^\infty\}) := 0,$$
$$P(((\tilde{A}_{i-1})^c \cap \Delta(x)) \times \Delta(1^i)) := P_X((\tilde{A}_{i-1})^c \cap \Delta(x)), \text{ and}$$
$$P(\tilde{A}_{i-1} \times \Delta(1^i)) := 0,$$

where 1^i is the string consisting of i 1s. For the construction of the measure, see Fig. 2. By construction, the total measure of P is one. Let $C := \bigcup_i B_i \times \Delta(1^i 0)$. For all $x, y \in S$, we have

$$P(x, y) = P((\Delta(x) \times \Delta(y)) \cap C) \text{ if } y \neq 1^{|y|},$$
$$P(x, y) = P_X(\Delta(x) \cap (\tilde{A}_{|y|-1})^c) \text{ if } y = 1^{|y|}.$$

Since the right-hand sides are computable, P is computable.

Next, we show (ii). Since $P(\Omega \times \Delta(1^k)) = P_X((\tilde{A}_{k-1})^c)$, we have $P_Y(\{1^\infty\}) = P(\Omega \times \Delta(1^\infty)) = P_X(\cap_k(\tilde{A}_k)^c) = 1 - \alpha > 0$ and $\{1^\infty\} \in \mathcal{R}^{P_Y}$. We have $P(x \mid 1^\infty) = 0$ if $\Delta(x) \subseteq \cup_i \tilde{A}_i$ and $P(x \mid 1^\infty) = P_X(x)/(1-\alpha)$ if $\Delta(x) \subseteq \cap_i(\tilde{A}_i)^c$. Since α is not computable, $P(\cdot \mid 1^\infty)$ is not computable.

By construction, we have $P_Y(\{1^i 0^\infty\}) > 0$ for all i and $P_Y(\{1^\infty\}) > 0$. Since $P_Y((\bigcup_i\{1^i 0^\infty\} \cup \{1^\infty\})^c) = 0$, the set $(\bigcup_i\{1^i 0^\infty\} \cup \{1^\infty\})^c$ is covered by a test, and we have $\mathcal{R}^{P_Y} = \bigcup_i\{1^i 0^\infty\} \cup \{1^\infty\}$. By construction, $P(\cdot \mid y^\infty)$ is orthogonal for different $y^\infty \in \mathcal{R}^{P_Y}$; we obtain statement (iii) of the theorem.

Similar to [4, 5], set $U_n := \{s \mid \exists i \ g(s) < g(a_i) + \frac{1}{n}\}$. Then $U = \{(s, n) \mid \exists n \ s \in U_n\}$ is r.e. and $P(\tilde{U}_n \mid 1^\infty) \leq 1/n$ for all n. U is a test that covers α, i.e. $\alpha \notin \mathcal{R}^{P(\cdot \mid 1^\infty)}$.

Let $b_1 \sqsubset b_2 \cdots$ be an increasing sequence of prefixes of α such that $b_n \to \alpha$ as $n \to \infty$. By construction, $\forall k \exists M \forall n \geq M \ P_{Y|X}(1^k \mid b_n) = 1$. We have $\forall k \ P_{Y|X}(1^k \mid \alpha) = 1$, $P_{Y|X}(\{1^\infty\} \mid \alpha) = 1$, and $\{1^\infty\} \in \mathcal{R}^{P_{Y|X}(\cdot \mid \alpha), \alpha}$. Since $\alpha \in \mathcal{R}^{P_X}$, from (1), we have $(\alpha, 1^\infty) \in \mathcal{R}^P$ and $\alpha \in \mathcal{R}^P_{1\infty}$. Since $\alpha \notin \mathcal{R}^{P(\cdot \mid 1^\infty)}$, from Theorem 3.8, we have the first part of statement (iv) of the theorem.

11

Since $\alpha \in \mathcal{R}_{1^\infty}^P$, $\alpha \notin \mathcal{R}^{P(\cdot|1^\infty)}$, $\mathcal{R}_{y^\infty}^P$ are disjoint for different $y^\infty \in \mathcal{R}^{P_Y}$ (Theorem 3.4), and $\mathcal{R}_{y^\infty}^P \supseteq \mathcal{R}^{P(\cdot|y^\infty)}$ for all $y^\infty \in \mathcal{R}^{P_Y}$ (Theorem 3.8), we have the latter part of statement (iv) of the theorem. $\qquad\square$

Proof of Theorem 2.4. Part 1 follows from

$$\mathcal{R}^{P_X} = \bigcup_{y^\infty \in \mathcal{R}^{P_Y}} \mathcal{R}_{y^\infty}^P \subseteq \bigcup_{y^\infty \in \mathcal{R}^{P_Y}} \bigcap_{y \to y^\infty} \mathcal{R}^{P(\cdot|y)} \tag{9}$$

$$\subseteq \bigcup_{y^\infty \in \mathcal{R}^{P_Y}} \mathcal{R}^{P_X} = \mathcal{R}^{P_X}, \tag{10}$$

where (9) follows from (4), and (10) follows from (7). By Lemma 3.7, we have $\mathcal{R}^{P_X} \supseteq \bigcup_{y^\infty \in \mathcal{R}^{P_Y}} \mathcal{R}^{P(\cdot|y^\infty)} \supseteq \bigcup_{y^\infty \in \mathcal{R}^{P_Y}} \mathcal{R}^{P(\cdot|y^\infty),y^\infty}$.

Part 2 follows from part 1 and Theorem 3.1.

Proof of part 3. The if part follows from part 1. The only if part follows from Theorem 3.4 (iii) and (5).

Proof of part 4. Let $P_Y(\{0^\infty\}) := 1$ and $P(\cdot \mid 0^\infty) := U$, where U is the uniform measure on X. Then $(X \times Y, \mathcal{B}_2, P)$ is computable and consistent. We have $P_X = U$ and $\mathcal{R}^{P_X} = \mathcal{R}^U = \bigcup_{y^\infty \in \mathcal{R}^{P_Y}} \mathcal{R}^{P(\cdot|y^\infty)} = \bigcup_{y^\infty \in \mathcal{R}^{P_Y}} \mathcal{R}^{P(\cdot|y^\infty),y^\infty} = \mathcal{R}^{P(\cdot|0^\infty)} = \mathcal{R}^{P(\cdot|0^\infty),0^\infty}$. Let $(X \times Y, \mathcal{B}_2, P')$ be the computable and consistent model defined in Theorem 3.10. Then $P_X = P'_X$ and $\bigcup_{y^\infty \in \mathcal{R}^{P_Y}} \mathcal{R}^{P(\cdot|y^\infty)} = \bigcup_{y^\infty \in \mathcal{R}^{P_Y}} \mathcal{R}^{P(\cdot|y^\infty),y^\infty} \neq \bigcup_{y^\infty \in \mathcal{R}'_Y} \mathcal{R}^{P'(\cdot|y^\infty)} = \bigcup_{y^\infty \in \mathcal{R}'_Y} \mathcal{R}^{P'(\cdot|y^\infty),y^\infty}$. We have the proof.

Proof of part 5. By definition of ML-random sequences, the sets (i), (iii), and (iv) are Borel sets. By definition of ML-random points, $\mathcal{R}^P \in \mathcal{B}_2$. By Fubini theorem, $\mathcal{R}_{y^\infty}^P \in \mathcal{B}_1$. By Theorem 3.1 and 3.8, the set (iv) is the smallest one among the sets (i)–(iv), and we have the first part of part 1. The latter part follows from part 1. $\qquad\square$

Remark 3.11 In Theorem 3.10, $P(\cdot \mid 1^\infty)$ is not computable with oracle 1^∞, however $\mathcal{R}^{P(\cdot|1^\infty)} = \mathcal{R}^{P(\cdot|1^\infty),1^\infty}$. The author does not know if the statement $\forall y^\infty \in \mathcal{R}^{P_Y}$ $\mathcal{R}^{P(\cdot|y^\infty)} = \mathcal{R}^{P(\cdot|y^\infty),y^\infty}$ is always true for the consistent Bayes models.

4. Discussions

We discuss the three topics: Bayes models and uniformly computable parametric models, an algorithmic solution to a classical statistical problem, and randomness in statistical models.

4.1. Bayes models and uniformly computable parametric models

Before we show our results, we summarize several known results for uniform tests [6, 13, 15, 21, 35]. Let $\{(\Omega, \mathcal{B}_1, P_{y^\infty}) \mid y^\infty \in \Omega\}$ be a parameterized family of distributions (parametric model $\{P_{y^\infty} \mid y^\infty \in \Omega\}$ for short). We assume that $y^\infty \to P_{y^\infty}(A)$ is measurable for any fixed $A \in \mathcal{B}_1$ [8] and uniformly computable [6, 13, 15, 21, 35]. If P_{y^∞} is uniformly

computable, then $P_{y^\infty}(\cdot)$ is defined for all y^∞ and the function $y^\infty \to P_{y^\infty}(x)$ is continuous for each fixed $x \in S$.

A function $t\colon \Omega \to [0,\infty]$ is called lower semicomputable if the set $\{x^\infty \mid r < t(x^\infty)\}$ is effectively open uniformly in rational r. A lower semicomputable function $t\colon \Omega \times \Omega \to [0,\infty]$ is defined in a similar manner [6]. We call a function $t_A\colon \Omega \to [0,\infty]$ a *blind test* w.r.t. P and oracle A if t_A is lower semicomputable with oracle A and $\int t_A(x^\infty)dP \le 1$. Let $T(P,A)$ be the set of blind tests w.r.t. P and oracle A. The set of blind random sequences w.r.t. P and oracle A is defined by $\mathrm{Blind}(P,A) := \{x^\infty \mid t_A(x^\infty) < \infty \text{ for all } t_A \in T(P,A)\}$. The set of blind random sequences w.r.t. P, $\mathrm{Blind}(P)$ is defined similarly. A lower semicomputable function $t\colon \Omega \times \Omega \to [0,\infty]$ is called a *uniform test* w.r.t. a parametric model $\{P_{y^\infty} \mid y^\infty \in \Omega\}$ if $\int t(x^\infty, y^\infty)dP_{y^\infty}(x^\infty) \le 1$ for all $y^\infty \in \Omega$.

The following proposition shows the relationship between $\mathcal{R}^P$ ($\mathcal{R}^{P,A}$), blind randomness, and randomness for uniformly computable parametric models.

Proposition 4.1 (Bienvenu et al. [6])
1. $\mathrm{Blind}(P) = \mathcal{R}^P$ and $\mathrm{Blind}(P,A) = \mathcal{R}^{P,A}$ for all probability measure P on $(\Omega, \mathcal{B})$ and oracle A. Assume that P are computable and computable with oracle A then $\mathcal{R}^P$ and $\mathcal{R}^{P,A}$ coincide with Martin-Löf random sequences and those with oracle A [23], respectively.
2. Assume that $\{P_{y^\infty} \mid y^\infty \in \Omega\}$ is uniformly computable. Then, there is a universal uniform test $\mathbf{t}$, i.e. for each uniform test t w.r.t. P_{y^∞} for $y^\infty \in \Omega$ there is $c > 0$ such that $t \le c\mathbf{t}$, and $\{x^\infty \mid \mathbf{t}(x^\infty, y^\infty) < \infty\} = \mathrm{Blind}(P_{y^\infty}, y^\infty) = \mathcal{R}^{P_{y^\infty}, y^\infty}$ for all $y^\infty \in \Omega$.

A Bayes model $(X \times Y, \mathcal{B}_2, P)$ is defined from a parametric model $\{P_{y^\infty} \mid y^\infty \in \Omega\}$ and a prior $(Y, \mathcal{B}_1, P_Y)$ by

$$P(x,y) := \int_{\Delta(y)} P_{y^\infty}(x)dP_Y(y^\infty) \text{ for all } x, y \in S. \tag{11}$$

Theorem 4.2 (Vovk and Y'yugin[35]) *Assume that a parametric model $\{P_{y^\infty} \mid y^\infty \in \Omega\}$ is uniformly computable and $(Y, \mathcal{B}_1, P_Y)$ is computable. Let $(X \times Y, \mathcal{B}_2, P)$ be the Bayes model defined by (11). Then*

$$\mathcal{R}^P_{y^\infty} = \mathcal{R}^{P_{y^\infty}, y^\infty} \text{ for all } y^\infty \in \mathcal{R}^{P_Y} \text{ and } \mathcal{R}^{P_X} = \bigcup_{y^\infty \in \mathcal{R}^{P_Y}} \mathcal{R}^P_{y^\infty}.$$

Definition 4.3 (Bienvenu et al. [6]) *A uniformly computable parametric model $\{P_{y^\infty} \mid y^\infty \in \Omega\}$ is called effectively orthogonal if $\mathcal{R}^{P_{y^\infty}, y^\infty} \cap \mathcal{R}^{P_{z^\infty}, z^\infty} = \emptyset$ for all $y^\infty \ne z^\infty$.*

Theorem 4.4 (Theorem 5.41 in Bienvenu et al. [6]) *Assume that a parametric model $\{P_{y^\infty} \mid y^\infty \in \Omega\}$ is uniformly computable and effectively orthogonal. Then*

$$\mathcal{R}^{P_{y^\infty}, y^\infty} = \mathcal{R}^{P_{y^\infty}} \text{ for all } y^\infty. \tag{12}$$

We state our results for uniform tests.

Definition 4.5 $P(\cdot \mid y^\infty)$ is called continuous at $y^\infty \in \mathcal{R}^{P_Y}$ if $\lim_n P(x \mid y^\infty(n)) = P(x \mid y^\infty)$ for all $x \in S$ and for any sequence $y^\infty(1), y^\infty(2), \ldots \in \mathcal{R}^{P_Y}$ such that $\lim_n y^\infty(n) \to y^\infty \in \mathcal{R}^{P_Y}$.

Theorem 4.6 Assume that a parametric model $\{P_{y^\infty} \mid y^\infty \in \Omega\}$ is uniformly computable and $(Y, \mathcal{B}_1, P_Y)$ is computable. Let $(X \times Y, \mathcal{B}_2, P)$ be the Bayes model defined by (11). Then,
1. P is computable. For all $y^\infty \in \mathcal{R}^{P_Y}$, $P(\cdot \mid y^\infty) = P_{y^\infty}(\cdot)$, $P(\cdot \mid y^\infty)$ is continuous and computable with oracle y^∞, $\mathcal{R}^{P_{y^\infty}, y^\infty} = \mathcal{R}^{P(\cdot \mid y^\infty), y^\infty}$, and $\mathcal{R}^{P_{y^\infty}} = \mathcal{R}^{P(\cdot \mid y^\infty)}$.
2. In addition, if $\{P_{y^\infty} \mid y^\infty \in \Omega\}$ is effectively orthogonal then, P is consistent.

Proof) Since P_{y^∞} is uniformly computable and P_Y is computable, P is computable. From (11), for all $x, y \in S$

$$P(x \mid y) = \int_{\Delta(y)} P_{y^\infty}(x) dP_Y(y^\infty) / P_Y(y). \tag{13}$$

Since P_{y^∞} is uniformly computable, $y^\infty \to P_{y^\infty}(x)$ is continuous for all $x \in S$. From (13), for all $x \in S$ and $y^\infty \in \Omega$,

$$\forall \epsilon > 0 \; \exists y \; \forall y' \; |P(x \mid y') - P_{y^\infty}(x)| < \epsilon \text{ for } y \sqsubseteq y' \sqsubset y^\infty.$$

By Theorem 2.1, for all $x \in S$ and $y^\infty \in \mathcal{R}^{P_Y}$

$$\forall \epsilon > 0 \; |P(x \mid y^\infty) - P_{y^\infty}(x)| < \epsilon.$$

Since $\epsilon > 0$ is arbitrary, we have that $P(\cdot \mid y^\infty) = P_{y^\infty}(\cdot)$ for all $y^\infty \in \mathcal{R}^{P_Y}$ and the latter statements in part 1. Continuity of $P(\cdot \mid y^\infty)$ follows from that of P_{y^∞}.
Proof of part 2. Assume that the parametric model is effectively orthogonal. By Theorem 4.2, $\mathcal{R}^P_{y^\infty} \cap \mathcal{R}^P_{z^\infty} = \emptyset$ for all $y^\infty, z^\infty \in \mathcal{R}^{P_Y}$ and $y^\infty \neq z^\infty$. If $y^\infty \notin \mathcal{R}^{P_Y}$, we have $\mathcal{R}^P_{y^\infty} = \emptyset$, see [20] or [31]. By Theorem 3.4, P is consistent. $\qquad\square$
The following proposition shows that neither of the converses of Theorem 4.6 is true.

Proposition 4.7 There is a computable and consistent $(X \times Y, \mathcal{B}_2, P)$ such that
(i) $P(\cdot \mid y^\infty)$ is computable with oracle y^∞ for all $y^\infty \in \mathcal{R}^{P_Y}$ and
(ii) $0^\infty \in \mathcal{R}^{P_Y}$ and $P(\cdot \mid y^\infty)$ is not continuous at 0^∞.

Proof) Let $r(0^n 1^\infty) := 2^{-n}$ for $n \geq 1$ and $r(0^\infty) := 1$. Let $P_{y^\infty}(x) := r(y^\infty)^{\sum x_i}(1 - r(y^\infty))^{|x| - \sum x_i}$ for all $x = x_1 \cdots x_{|x|} \in S, x_i \in \{0, 1\}$ if $y^\infty = 0^n 1^\infty$ for $n \geq 1$. Let $P_{y^\infty}(x) := 1$ for all $x \sqsubset 1^\infty$ if $y^\infty \neq 0^n 1^\infty$ for $n \geq 1$. Let $P_Y(\{0^n 1^\infty\}) := 2^{-(n+1)}$ and $P_Y(0^\infty) := 1/2$. Then P_Y is a computable probability measure and $\mathcal{R}^{P_Y} = \bigcup_{n \geq 1} \{0^n 1^\infty\} \cup \{0^\infty\}$. $P_{y^\infty} \perp P_{z^\infty}$ if $y^\infty, z^\infty \in \mathcal{R}^{P_Y}$ and $y^\infty \neq z^\infty$. Let P be a Bayes model defined by (11). Then, P is computable and $P(\cdot \mid y^\infty) = P_{y^\infty}$ for all $y^\infty \in \mathcal{R}^{P_Y}$. By Theorem 3.4, P is consistent. Since $P(1 \mid 0^n 1^\infty) = 2^{-n}$ and $P(1 \mid 0^\infty) = 1$, $P(\cdot \mid y^\infty)$ is not continuous at $0^\infty \in \mathcal{R}^{P_Y}$. $\qquad\square$
By Theorem 4.6 and Proposition 4.7, the assumption of Theorem 3.1 and 3.2 is weaker than that of Theorem 4.2. In [35], a quantitative version of Theorem 4.2 is shown, for more details, see Section 7 in [5] and [26]. The equation (6) in Theorem 3.8 is true under

computable consistent Bayes models. By Theorem 4.6 and Proposition 4.7, those models constitute a larger family of models than those defined by (11) with uniformly computable and effectively orthogonal parametric models and computable priors. Note that the equality in (12) is true for all $y^\infty \in \Omega$ while that in (6) is true for all $y^\infty \in \mathcal{R}^{P_Y}$. In [6], Theorem 4.4 is proved for the families of measures that are parameterized by themselves. Kjos Hanssen [15] demonstrated a different proof of (12) for the Bernoulli model.

Definition 4.8 *A measurable function (estimator) $f\colon X \to Y$ is called consistent for the parametric model $\{P_{y^\infty} \mid y^\infty \in \Omega\}$ if $P_{y^\infty}(f^{-1}(y^\infty)) = 1$ for all $y^\infty \in \Omega$. A parametric model $\{P_{y^\infty} \mid y^\infty \in \Omega\}$ is called consistent if there exists a consistent estimator.*

In Definition 8.1 [6], uniformly computable consistent models are called orthogonal. We use the term consistent for the orthogonal models to distinguish pairwise orthogonality from orthogonality.

Theorem 4.9 *Assume that $\{P_{y^\infty} \mid y^\infty \in \Omega\}$ is uniformly computable. Let*

$$f^{-1}(y^\infty) := \begin{cases} \mathcal{R}^{P_{y^\infty}, y^\infty} & \text{if } y^\infty \neq 0^\infty, \\ \mathcal{R}^{P_{0^\infty}, 0^\infty} \cup (\Omega \setminus \bigcup_{y^\infty \in \Omega} \mathcal{R}^{P_{y^\infty}, y^\infty}) & \text{if } y^\infty = 0^\infty. \end{cases}$$

The following statements are equivalent.
(i) $\{P_{y^\infty} \mid y^\infty \in \Omega\}$ is effectively orthogonal.
(ii) $f\colon \Omega \to \Omega$ is well-defined and a consistent estimator for $\{P_{y^\infty} \mid y^\infty \in \Omega\}$.

We give the proof after Definition 4.10, Proposition 4.11, and Lemma 4.12.

Definition 4.10 *Let A be a subset of $X \times Y$ and B be a subset of Y. Set*

$$\pi_X(A) := \{x^\infty \mid \exists y^\infty \in Y \ (x^\infty, y^\infty) \in A\} \ and \tag{14}$$
$$\pi_{X,B}(A) := \{x^\infty \mid \exists y^\infty \in B \ (x^\infty, y^\infty) \in A\}.$$

A countable union of closed sets is called F_σ set.

Proposition 4.11 (a special case of Exercise 2.3.24 [29] pp.61) *Assume that $C \subseteq X \times Y$ is closed. Then $\pi_X(C)$ is closed.*

Proof) Let $\alpha(1), \alpha(2), \ldots$ be a sequence of points in $\pi_X(C)$. Suppose that $\alpha(i) \to \alpha$ as $i \to \infty$. We show that $\alpha \in \pi_X(C)$. By (14), there is a sequence $(\alpha(1), \beta(1)), (\alpha(2), \beta(2)), \ldots \in C$. Since $X \times Y$ is compact, C is compact and there is a subsequence $(\alpha(n(1)), \beta(n(1))), (\alpha(n(2)), \beta(n(2))), \ldots \in C$ such that $n(1) < n(2) < \cdots$ and $(\alpha(n(i)), \beta(n(i))) \to (\alpha', \beta') \in C$ as $i \to \infty$. Thus $\lim_i \alpha(i) = \lim_i \alpha(n(i)) = \alpha' = \alpha \in \pi_X(C)$. □

Lemma 4.12 *Assume that A is F_σ in $X \times Y$. Then $\pi_X(A)$ is F_σ. Further assume that $\pi_{X,C}(A) \cap \pi_{X,D}(A) = \emptyset$ for all $C \cap D = \emptyset$. Then $\pi_{X,C}(A) \in \mathcal{B}_1$ for all $C \in \mathcal{B}_1$.*

Proof) Since A is F_σ, there are closed sets A_n in $X \times Y$ for all n and $A = \bigcup_n A_n$. First we show that $\pi_X(A) \in \mathcal{B}_1$. We have

$$\pi_X(A) = \bigcup_{y^\infty \in Y} A_{y^\infty} = \bigcup_{y^\infty \in Y} \bigcup_n (A_n)_{y^\infty} = \bigcup_n \bigcup_{y^\infty \in Y} (A_n)_{y^\infty} = \bigcup_n \pi_X(A_n).$$

By Proposition 4.11, $\pi_X(A_n)$ is closed and $\pi_X(A)$ is F_σ. Similarly,
1. $\pi_{X,\Delta(s)}(A)$ is F_σ for all $s \in S$.
2. Assume that $\pi_{X,C_k}(A) \in \mathcal{B}_1$ for all k. Then $\pi_{X,\bigcup_k C_k}(A) = \bigcup_k \pi_{X,C_k}(A) \in \mathcal{B}_1$.
3. Assume that $\pi_{X,C}(A) \in \mathcal{B}_1$. By assumption, $\pi_{X,C^c}(A) = \pi_X(A) \setminus \pi_{X,C}(A) \in \mathcal{B}_1$.
Let $\mathcal{B}' := \{C \mid \pi_{X,C}(A) \in \mathcal{B}_1\}$. From 1–3, $\{\Delta(s) \mid s \in S\} \subseteq \mathcal{B}'$ and $\mathcal{B}'$ is a σ-algebra. By definition, we have $\mathcal{B}_1 \subseteq \mathcal{B}'$. $\qquad\square$

Proof of Theorem 4.9) First we prove (i) $\Rightarrow$ (ii). Let $\mathbf{t}$ be a universal uniform test and $A_n := \{(x^\infty, y^\infty) \mid \mathbf{t}(x^\infty, y^\infty) > n\}$ for all n. Then A_n is open, $\mathcal{R}^{P_{y^\infty}, y^\infty} = ((\cap_n A_n)^c)_{y^\infty}$ for all $y^\infty \in \Omega$, and $A := (\cap_n A_n)^c$ is a F_σ set. Assume (i). Then $f \colon \Omega \to \Omega$ is well-defined, i.e. $f^{-1}(y^\infty) \cap f^{-1}(z^\infty) = \emptyset$ for all $y^\infty \neq z^\infty$. Since the set A satisfies the assumptions of Lemma 4.12, f is measurable. Thus f is consistent and we have (ii). The implication (ii) $\Rightarrow$ (i) is obvious. $\qquad\square$

In [8], necessary and sufficient conditions for consistency of parametric models are shown for general parametric models.

4.2. Algorithmic solution to a classical statistical problem

In parametric models, estimators that are consistent at all parameters are concerned, while in Bayes models, those that are consistent at almost all parameters are concerned. The identification of the points at which the posterior distributions weakly converge constitutes the problem (see pp.4 Diaconis and Freedman [9] and pp.24 Ghosh and Ramamoorthi [14]).

Let $(X \times Y, \mathcal{B}_2, P)$ be computable, where $X = Y = \Omega$. By Theorem 3.4, we have an algorithmic solution to the problem, i.e. the Bayes model P is consistent if and only if

$$\mathcal{R}^{P_Y} \subseteq \{y^\infty \mid \text{the posterior distribution is consistent at } y^\infty\}. \tag{15}$$

Remark 4.13 1. Assume that the Bayes model is computable and consistent. Then, by (15), we know that the posterior distribution is always consistent at the ML-random parameters without explicitly computing the posterior distribution. In other words, if the Bayes model is consistent and computable and the posterior distribution is not consistent at some parameters, then these parameters are not ML-random. Freedman [12] and Schwartz [27] identified the sets of the consistent parameters for smooth finite dimensional i.i.d. (independent and identically distributed) models, see pp.4 [9] or the examples below. Though (15) identifies a subset of consistent parameters, Theorem 3.4 and (15) give a new solution to the problem since they even hold true for any consistent P on $(\Omega^2, \mathcal{B}_2)$ by relativizing with oracle P without any further assumption.

2. It is straightforward to extend the equivalence of the statements except for (viii) in Theorem 3.4 to joint probabilities on complete separable metric spaces such as $\mathbb{R}^\infty \times \mathbb{R}^\infty$, see Doob [11] or pp.24 Remark 1.3.3 [14]. For example, let $(\mathbb{R} \times \mathbb{R}, \mathcal{B})$ be a measurable space where $\mathcal{B}$ is the Borel σ-algebra. Then, Theorem 3.4 holds for computable probabilities on $(\mathbb{R} \times \mathbb{R}, \mathcal{B})$ by replacing open intervals $\Delta(s)$, $s \in S$ in statements (iii) and (iv) in Theorem 3.4 with half open intervals $[a, b)$, $-\infty \le a < b \le \infty$, $a, b \in \mathbb{Q} \cup \{-\infty, \infty\}$. Similarly, we can extend Theorem 3.4 to probability measures on $\mathbb{R}^\infty \times \mathbb{R}^\infty$.

We examine (15) with examples of consistent Bayes models on complete separable metric spaces.

Example 4.14 (Example 2 pp.17 in Schwartz [27]) Let $\Theta = [1, 2)$ be the parameter space and the prior P_Θ the uniform measure on Θ. Let $X_1, X_2, \ldots$ be i.i.d. random variables and $X^n := (X_1, \ldots, X_n)$ where X_1 obeys the uniform measure on $[0, 1)$ if $\theta = 1$ and on $[0, 2/\theta)$ if $1 < \theta < 2$. Let $p_\theta(X^n)$ be the probability density function for $\theta \in \Theta$. Then $p_\theta(X^n) = 1$ if $\theta = 1$ else $(\theta/2)^n$. Let $\hat{\theta}(X^n) := \arg\max_\theta p_\theta(X^n)$ and $Y_n := \max_{1 \le i \le n} X_i$. Then $\hat{\theta} = 1$ if $Y_n \le 1$ else $2/Y_n$. $\hat{\theta}$ is consistent for all parameters. The Bayes model is computable. By Theorem 3.4, the Bayes model is consistent, and we have (15). The expectation of the parameter w.r.t. the posterior distribution (Bayes estimate) is

$$\beta_n(X^n) = \begin{cases} \frac{n+1}{n+2} \frac{2^{n+2}-1}{2^{n+1}-1} & \text{if } Y_n \le 1 \\ \frac{\int_1^{2/Y_n} \theta^{n+1} d\theta}{\int_1^{2/Y_n} \theta^n d\theta} & \text{else,} \end{cases}$$

and the posterior distribution is consistent except for $\theta = 1$. Since P_Θ is the uniform measure, $1 \notin \mathcal{R}^{P_\Theta}$, and (15) is true. Let $P'_\Theta := \alpha P_\Theta + (1 - \alpha)\delta_1$, where $0 < \alpha < 1$ and $\delta_1(\{1\}) = 1$. Then, the posterior distribution with prior P'_Θ is consistent for all θ. Since $P'_\Theta(\{1\}) > 0$, $1 \in \mathcal{R}^{P'_\Theta}$, and (15) is true.

Before we show Example 4.19, we prove Proposition 4.17 for constructive topological spaces.

Definition 4.15 (Constructive topological space [6, 13, 24]) *Let $(X, \mathcal{T})$ be a topological space with countable base $\mathcal{O}$, i.e. every open set $A \in \mathcal{T}$ is a union of elements of $\mathcal{O}$. Let D be a subset of $\mathbb{N}$ and $f \colon D \to \mathcal{O}$ an onto function (naming system p.132 [13]). We call $(X, \mathcal{T}, \mathcal{O}, f, D)$ constructive topological space if D is r.e. The countable base $\mathcal{O}$ is called r.e. if $(X, \mathcal{T}, \mathcal{O}, f, D)$ is a constructive topological space.*

Definition 4.16 (ML-randomness on constructive topological space) *Let $(X, \mathcal{T}, \mathcal{O}, f, D)$ be a constructive topological space and $(X, \mathcal{B}, P)$ a probability space where $\mathcal{B}$ is the σ-algebra generated by $\mathcal{O}$. Then ML-test and $\mathcal{R}^P$ w.r.t. P are defined with similar manner in Section 2, i.e. $U \subseteq D \times \mathbb{N}$ is called ML-test or test w.r.t. P if U is r.e. and for all n $\tilde{U}_n \supseteq \tilde{U}_{n+1}$ and $P(\tilde{U}_n) < 2^{-n}$ where $U_n := \{x \mid (x, n) \in U\}$ and $\tilde{A} := \bigcup_{s \in A} f(s)$ for $A \subseteq D$. Then $\mathcal{R}^P := (\bigcup_{U \colon test} \bigcap_n \tilde{U}_n)^c$.*

Proposition 4.17 *Assume that $(Y, \mathcal{T}, \mathcal{O}, f, D)$ is a constructive topological space. Let $(Y, \mathcal{B}, P)$ be a probability measure where $\mathcal{B}$ is the σ-algebra generated by $\mathcal{O}$. Then*

$$\mathcal{R}^P \subseteq \text{ the support of } P, \tag{16}$$

where the support of P is the complement of the largest open null set w.r.t. P.

Proof) The complement of the support is a finite or countable union of null sets of $\mathcal{O}$. For each null set $A \in \mathcal{O}$, there is i such that $f(i) = A$ and $P(A) = 0$, i.e. A is covered by a test. Since a countable union of tests covers the complement of the support, we have the proposition. $\qquad\qquad\square$

For simplicity, we write X for a constructive topological space $(X, \mathcal{T}, \mathcal{O}, f, D)$.

Remark 4.18 1. $\mathcal{R}^P$ in (16) is blind randomness since we do not assume that P is computable in Definition 4.16.
2. $\{(a, b) \mid a < b, a, b \in \mathbb{Q}\}$ is an r.e base for $\mathbb{R}$ with usual topology. Similarly, for a positive integer m, $\mathbb{R}^m$ and $\mathbb{R}^\infty$ with usual topologies have r.e. bases and are constructive topological spaces. For other examples of r.e. bases and constructive topological spaces, see [13].

Example 4.19 (Freedman [12] Schwartz [27]) Freedman [12] and Schwartz [27] showed that for smooth i.i.d. parametric models with finite-dimensional parameter space $\mathbb{R}^m$, posterior distributions are consistent on the support and inconsistent outside the support. Since $\mathbb{R}^m$ has an r.e. base, by (16), we have (15) for these models.

4.3. Randomness in statistical models

We compare randomness in Bayes models with that in non-Bayes models.

Example 4.20 (Uniformly computable parametric models) Assume that $\{P_{y^\infty} \mid y^\infty \in \Omega\}$ is uniformly computable and effectively orthogonal. Let $g^{-1}(y^\infty) := \mathcal{R}^{P_{y^\infty}, y^\infty}$ for all $y^\infty \in \Omega$. Then $g \colon \bigcup_{y^\infty \in \Omega} \mathcal{R}^{P_{y^\infty}, y^\infty} \to \Omega$ is well-defined and surjective. By Lemma 4.12, g is measurable. We have $g(x^\infty) = y^\infty$ and $x^\infty \in \mathcal{R}^{P_{y^\infty}, y^\infty}$ for all $x^\infty \in \bigcup_{y^\infty \in \Omega} \mathcal{R}^{P_{y^\infty}, y^\infty}$.

In the example above, first we consider the set of random sequences $\bigcup_{y^\infty \in \Omega} \mathcal{R}^{P_{y^\infty}, y^\infty}$, the union of the sets of random sequences w.r.t. the parameterized family of models. Then for uniformly computable and effectively orthogonal models, for all $x^\infty \in \bigcup_{y^\infty \in \Omega} \mathcal{R}^{P_{y^\infty}, y^\infty}$, the true parameter y^∞ is estimated by $g(x^\infty)$ $(g(x^\infty) = y^\infty)$ and x^∞ is random w.r.t. P_{y^∞} $(x^\infty \in \mathcal{R}^{P_{y^\infty}, y^\infty})$. For a similar argument on collectives, see von Mises [25].

In Bayes models, we assume that a sequence x^∞ is random w.r.t. the marginal distribution on sample space. Assumption (i') requires that x^∞ is also a member of the union of the sets of random sequences w.r.t. conditional distributions given random parameters. Then for consistent models, for all random sequence x^∞ w.r.t. the marginal distribution on sample space, the true parameter y^∞ is estimated by the posterior distributions given finite prefixes of x^∞, x^∞ is random w.r.t. the conditional distribution given y^∞, and y^∞ is random w.r.t. the prior. In other words, we may say that our notion of randomness is a Bayes version of that in uniformly computable parametric models.

Acknowledgment

References

[1] Ackerman, N.L., Freer, C.E., Roy, D.M., 2011. On the computability of conditional probability, in: IEEE 26th Annual Symposium on Logic in Computer Science, pp. 107–116. Arxiv:1005.3014.

[2] Barron, A., Rissanen, J., Yu, B., 1998. The minimum description length principle in coding and modeling. IEEE Trans. Inform. Theory 44, 2743–2760.

[3] Barron, A.R., 1985. Logically smooth density estimation. Ph.D. thesis. Stanford Univ.

[4] Bauwens, B., 2017. Conditional measure and the violation of van Lambalgen's theorem for Martin-Löf randomness. Theory Comput. Syst. 60, 314–323. Arxiv:1103.1529.

[5] Bauwens, B., Shen, A., Takahashi, H., 2017. Conditional probabilities and van Lambalgen theorem revisited. Theory Comput. Syst. 61, 1315–1336.

[6] Bienvenu, L., Gács, P., Hoyrup, M., Rojas, C., Shen, A., 2011. Algorithmic tests and randomness with respect to a class of measures. Proc. of the Steklov Institute of Mathematics 274, 41–102. Arxiv:1103.1529v2.

[7] Billingsley, P., 1995. Probability and Measures. 3rd ed., Wiley.

[8] Breiman, L., LeCam, L., Schwartz, L., 1964. Consistent estimates and zero-one sets. Ann. Math. Statist. 35, 157–161.

[9] Diaconis, P., Freedman, D., 1986. On the consistency of Bayes estimates. Ann. Statist. 14, 1–26.

[10] Dębowski, Ł., 2009. Computable Bayesian compression for uniformly discretizable statistical models, in: ALT 2009, pp. 53–67.

[11] Doob, J.L., 1948. Application of the theory of martingales, in: Le Calcul des Probabilités et ses Applications. Colloq. Intern. du C.N.R.S., Paris. pp. 22–28. The paper is available from B. Locker, Doob at Lyon, www.jehps.net.

[12] Freedman, D.A., 1963. On the asymptotic behavior of Bayes' estimates in the discrete case. Ann. Math. Statist. 34, 1386–1403.

[13] Gács, P., 2005. Uniform test of algorithmic randomness over a general space. Theoret. Comput. Sci. 341, 91–137.

[14] Ghosh, J.K., Ramamoorthi, R.V., 2003. Bayesian Nonparametrics. Springer.

[15] Hanssen, B.K., 2010. The probability distribution as a computational resource for randomness testing. Journal of Logic and Analysis 2, 1–13.

[16] Kolmogorov, A.N., 1933. Grundbegriffe der Wahrscheinlichkeitsrechnung. volume 2 of *Eng. Math.* Springer Verlag, Berlin.

[17] Kolmogorov, A.N., 1963. On tables of random numbers. Sankhyā 25, 369—376.

[18] Kolmogorov, A.N., 1965. Three approaches to the quantitative definition of information. Probl. Inf. Transm. 1, 1–7.

[19] Kolmogorov, A.N., 1968. Logical basis for information theory and probability theory. IEEE Trans. Inform. Theory 14, 662–664.

[20] van Lambalgen, M., 1987. Random sequences. Ph.D. thesis. Universiteit van Amsterdam.

[21] Levin, L.A., 1976. Uniform tests of randomness. Soviet Math. Dokl. 17, 337–340.

[22] Li, M., Vitányi, P., 2019. An introduction to Kolmogorov complexity and Its applications. Forth ed., Springer, New York.

[23] Martin-Löf, P., 1966. The definition of random sequences. Information and Control 9, 602–609.

[24] Martin-Löf, P., 1968. Notes on constructive mathematics. Almqvist & Wiksell, Stockholm.

[25] von Mises, R., 1981. Probability, Statistics and Truth. Dover.

[26] Romashchenko, A., Shen, A., Zimand, M., 2021. 27 open problems in Kolmogorov complexity. ACM SIGACT News 52, 31–54.

[27] Schwartz, L., 1965. On Bayes procedures. Z. Wahrscheinlichkeitstheorie 4, 10–26.

[28] Shen, A., Uspensky, V.A., Vereshchagin, N.K., 2017. Kolmogorov complexity and algorithmic randomness. AMS.

[29] Srivastava, S.M., 1998. A course on Borel sets. Springer.

[30] Takahashi, H., 2006. Bayesian approach to a definition of random sequences and its applications to statistical inference, in: 2006 IEEE International Symposium on Information Theory, pp. 2180–2184.

[31] Takahashi, H., 2008. On a definition of random sequences with respect to conditional probability. Inform. and Compt. 206, 1375–1382.

[32] Takahashi, H., 2009. Some problems of algorithmic randomness on product space, in: The 8th Workshop on Stochastic Numerics, RIMS Kôkyûroku, Kyoto University. pp. 175–196.

[33] Takahashi, H., 2011. Algorithmic randomness and monotone complexity on product space. Inform. and Compt. 209, 183–197.

[34] Takahashi, H., 2014. Generalization of van lambalgen's theorem and blind randomness for conditional probabilities. Arxiv:1310.0709.

[35] Vovk, V.G., V'yugin, V.V., 1993. On the empirical validity of the Bayesian method. J. R. Stat. Soc. B 55, 253–266.

[36] Williams, D., 1991. Probability with Martingales. Cambridge University Press, Cambridge.